Swiss Orthopaedics Minimal Dataset: First Pilot Report of Reliability and Validity

Jentzsch, T ; Dora, C ; Müller, U ; Farshad, M

Abstract: Background. The Swiss Orthopaedics Minimal Dataset (SOMD) was launched seven years ago. It is a standardized, generic, and patient-reported outcome questionnaire, comprising ten items (location of disease, pain within the past four weeks, limitations at work/leisure/sleep/autonomy, subjective value of a body part, employment status, work disability (sick leave/pension), and household support). We conducted this study about the SOMD to report its reliability, validity, and clinical applicability. Methods. A retrospective observational cohort study was conducted. The test-retest study population (n = 60; lost to follow-up: n = 7 (12%)) was drawn from three retirement homes (in 2013), while the test study population (n = 14,180; excluded (e.g., duplicates): n = 1,990 (14%)) consisted of patients from a university hospital (in 2014–2017). In the test-retest study population, the same questionnaire was completed twice (at days 0 and 7). In the test study population, only the first questionnaire was included (to avoid duplicates). In a subgroup of the test study population (n = 302), only those patients who completed the SOMD and Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) of the hip within 14 days were considered (to minimize recall bias). Reliability (test-retest and internal consistency), criterion validity for the item of pain, and return rates were analyzed. Results. The test-retest study population (n = 53) showed very high test-retest reliability for all tested items of the SOMD (intraclass correlation coefficient = 0.96–1.00 (95% confidence interval 0.93–1.00),). The test study population (n = 12,190) revealed good internal consistency reliability for all ten items (Cronbach's alpha = 0.80). The return rates of the SOMD were improvable (43% in 2016 and 31% in 2017). The subgroup of the test study population (n = 302) displayed a borderline acceptable criterion validity (correlation of the item of pain between SOMD and WOMAC hip: rho = 0.62,). Conclusion. This is the first report about the validation of the SOMD. A relatively high reliability (test-retest and internal consistency), borderline acceptable (criterion) validity for the item of pain, and improvable clinical implementation were observed. This analysis serves as the basis for a structured modification of the SOMD to improve its value.

DOI: https://doi.org/10.1155/2020/6673175

Originally published at:
Jentzsch, T; Dora, C; Müller, U; Farshad, M (2020). Swiss Orthopaedics Minimal Dataset: First Pilot Report of Reliability and Validity. Advances in Orthopedics, 2020:6673175.
DOI: https://doi.org/10.1155/2020/6673175

Hindawi
Advances in Orthopedics
Volume 2020, Article ID 6673175, 4 pages
https://doi.org/10.1155/2020/6673175

Research Article

Swiss Orthopaedics Minimal Dataset: First Pilot Report of Reliability and Validity

Thorsten Jentzsch,[1] **Claudio Dora**,[2] **Urs Müller**,[3] **and Mazda Farshad**[1]

[1]*Department of Orthopedics, Balgrist University Hospital, University of Zurich, Zurich, Switzerland*
[2]*Swiss Orthopaedics, Bourg-en-Lavaux, Switzerland*
[3]*Onaxis Ltd, Bern, Switzerland*

Correspondence should be addressed to Thorsten Jentzsch; thorsten.jentzsch@gmail.com

Received 18 October 2020; Accepted 10 December 2020; Published 30 December 2020

Academic Editor: Andreas K. Demetriades

Background. The Swiss Orthopaedics Minimal Dataset (SOMD) was launched seven years ago. It is a standardized, generic, and patient-reported outcome questionnaire, comprising ten items (location of disease, pain within the past four weeks, limitations at work/leisure/sleep/autonomy, subjective value of a body part, employment status, work disability (sick leave/pension), and household support). We conducted this study about the SOMD to report its reliability, validity, and clinical applicability. *Methods*. A retrospective observational cohort study was conducted. The test-retest study population ($n = 60$; lost to follow-up: $n = 7$ (12%)) was drawn from three retirement homes (in 2013), while the test study population ($n = 14{,}180$; excluded (e.g., duplicates): $n = 1{,}990$ (14%)) consisted of patients from a university hospital (in 2014–2017). In the test-retest study population, the same questionnaire was completed twice (at days 0 and 7). In the test study population, only the first questionnaire was included (to avoid duplicates). In a subgroup of the test study population ($n = 302$), only those patients who completed the SOMD and Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) of the hip within 14 days were considered (to minimize recall bias). Reliability (test-retest and internal consistency), criterion validity for the item of pain, and return rates were analyzed. *Results*. The test-retest study population ($n = 53$) showed very high test-retest reliability for all tested items of the SOMD (intraclass correlation coefficient = 0.96–1.00 (95% confidence interval 0.93–1.00), $p < 0.001$). The test study population ($n = 12{,}190$) revealed good internal consistency reliability for all ten items (Cronbach's alpha = 0.80). The return rates of the SOMD were improvable (43% in 2016 and 31% in 2017). The subgroup of the test study population ($n = 302$) displayed a borderline acceptable criterion validity (correlation of the item of pain between SOMD and WOMAC hip: rho = 0.62, $p < 0.001$). *Conclusion*. This is the first report about the validation of the SOMD. A relatively high reliability (test-retest and internal consistency), borderline acceptable (criterion) validity for the item of pain, and improvable clinical implementation were observed. This analysis serves as the basis for a structured modification of the SOMD to improve its value.

1. Introduction

Health-related quality of life (HRQoL) describes the perceived well-being of individuals [1]. This can change over time (e.g., after treatment of disease). Accurate measurements (e.g., questionnaires) to detect this change are important not only for internal and external quality control but also for the evaluation of treatment success and costs. There are many questionnaires to choose from, but the gold standard remains elusive.

The Swiss Orthopaedics Minimal Dataset (SOMD) was introduced seven years ago, in 2013, as a measuring tool for HRQoL in order to assess the indications and results of all orthopaedic surgeries. It is a standardized, generic, and patient-reported outcome questionnaire. It consists of ten items: location of disease (only at one body part), pain within the past four weeks (score of 0–100), limitations at work/leisure/sleep/autonomy (0–100 each), subjective value of a body part (0–100), employment status (training, employed, and retired), work disability ((A) sick leave and (B) pension)

(0–100), and household support (0, <1, 1–2, 3–5, and >5 hours per day and a care assistant). The interpretation is based on a rating scale for each item without a single index value. It is free-of-charge, available in four languages (German, English, French, and Italian), accessible in electronic- and paper-based forms, and can be completed within five minutes. In our institution, it is used at first consultation and intermediate and final follow-ups.

The psychometric properties (i.e., reliability, validity, and sensitivity) of the SOMD have not been adequately provided yet. So far, only content validity (i.e., expert opinion) has been established. The reliability, criterion validity, and clinical implementation (i.e., return rate) remain unknown. Reliability is defined as the consistency of an item. It can be assessed with a test-retest method providing an intraclass correlation coefficient (ICC) and internal consistency across different items within a test providing Cronbach's alpha (α). Criterion validity describes the correlation between an item of a new questionnaire with the same item of an established questionnaire providing Spearman's rho. A well-established, reliable, and valid questionnaire for comparison with the SOMD is the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), which was first described by Bellamy in a Master's thesis in 1982 [2–7]. It is one of the most commonly used questionnaires for the hip, available in around 100 languages, and consists of 24 questions about pain, stiffness, and daily activities. The item of pain refers to the last two days and different situations (walking, stairs, bed, sitting, and standing), which is particularly detail-oriented. Scoring is performed on a five-point (or level) Likert scale from none, over, and moderate to extreme.

We conducted the first study about the validation of the SOMD. The study hypothesis was that the SOMD is highly reliable, valid, and clinically applicable.

2. Methods

A retrospective observational cohort study was conducted. The cantonal ethics committee issued a waiver to allow this study with anonymous data without the need for informed consent (BASEC Request-Nr. 2018-00276).

The test-retest study population ($n = 60$) was drawn from retirement homes in 2013. The loss to follow-up was acceptable ($n = 7$ (12%)). The retirement homes offered assisted but autonomous living. After thorough instruction of individuals, the German version of the SOMD was completed twice, initially at day 0 and again at day 7. Two items, employment status and work disability (sick leave/pension), were not evaluated in this elderly study population.

The test study population ($n = 14{,}180$) consisted of all patients that filled out the questionnaire from a university hospital from April 2014 until December 2017. Only the first SOMD was considered. Duplicates and test questionnaires were excluded ($n = 1{,}990$ (14%)). Furthermore, a subgroup of the test study population was chosen, in which the WOMAC of the hip had also been completed within 14 days (to minimize recall bias) ($n = 302$).

Data were given as medians (interquartile range (IQR)). For the test-retest study population, the test-retest reliability (ICC (95% confidence interval (CI))) was calculated. For the test study population, the internal consistency (Cronbach's α of all ten items and item-rest correlations for each item), criterion validity (Spearman's correlation (rho) of the item of pain in the SOMD and WOMAC of the hip), and return rates were calculated. A scatterplot is provided for illustration. The significance level was set at $p < 0.05$. For a test-retest reliability test, it was previously suggested that a sample size of ≥46 individuals would be needed to measure an ICC of 0.9 with a power of 0.8 at a significance level of 5% [8]. For internal consistency, it was also suggested that a sample size of ≥300 individuals would be sufficient [9]. Both of our study populations surpassed these numbers. Stata (IC13.1; StataCorp, College Station, Texas, United States of America) was used for analysis.

3. Results

In the test-retest study population ($n = 53$), the median age was 75 (IQR 69–79) years. There were more females than males (females: $n = 34$ (57%) vs males: $n = 26$ (43%)). The scores for each item were moderate (e.g., pain: median = 50 (IQR 30–90)).

The test-retest reliability for all tested separate items of the SOMD was very high (e.g., pain: ICC = 1.00 (95% CI 1.00–1.00), $p < 0.001$) (Table 1). The best reliabilities were found for the items of pain, limitation at leisure and autonomy, subjective value of a body part, and household support (ICC = 1.00, each with CIs between 0.97 and 1.00).

In the test study population ($n = 12{,}190$), the median age was 49 years (IQR 35–61). There were less females than males (females: $n = 5{,}872$ (48%) and males: $n = 6{,}318$ (52%)). The scores for each item were moderate (e.g., pain: median = 50 (IQR 30–70)) (Table 2). The right side was most commonly affected (right: $n = 6{,}186$ (51%); left: $n = 4{,}615$ (38%); axial: $n = 1{,}376$ (11%); not applicable: $n = 13$ (0%)). The shoulder, knee, and foot were most commonly impaired (shoulder: $n = 2{,}390$ (20%); knee: $n = 2{,}348$ (20%); foot: $n = 1{,}969$ (17%); pelvis: $n = 1{,}393$ (11%); axial: $n = 1{,}376$ (11%); hand: $n = 664$ (5%); ankle: $n = 411$ (3%); hip: $n = 339$ (3%); elbow: $n = 253$ (2%); thigh: $n = 228$ (2%); calf: $n = 204$ (2%); wrist: $n = 359$ (3%); upper arm: $n = 156$ (1%); lower arm: $n = 87$ (0%); not applicable: 13 (0%)).

The internal consistency for all ten items was very high (Cronbach's $\alpha = 0.80$). The item-rest correlations were the lowest for employment status (0.18), work disability (sick leave (0.40), pension (0.24)), and household support (0.36). After the removal of these items from the calculation of the internal consistency, Cronbach's α increased to 0.85. The return rates of the SOMD were improvable (43% in 2016 and 31% in 2017).

Table 1: Descriptive data and test-retest reliability for the test-retest sample ($n = 60$).

Item	Test ($n = 60$) Median (IQR)	Retest ($n = 53$) Median (IQR)	ICC (95% CI)	P value*
Pain	50 (30–90)	60 (40–90)	1.00 (1.00–1.00)	<0.001
Work limitation	80 (25–100)	80 (40–100)	0.99 (0.99–1.00)	<0.001
Leisure limitation	50 (15–100)	60 (20–100)	1.00 (1.00–1.00)	<0.001
Sleep limitation	90 (25–100)	90 (40–100)	0.96 (0.93–0.98)	<0.001
Independence limitation	90 (25–100)	90 (50–100)	1.00 (0.97–0.99)	<0.001
Subjective body part value	70 (50–90)	70 (50–90)	1.00 (1.00–1.00)	<0.001

*F-test. IQR: interquartile range; ICC: intraclass correlation coefficient; %: percent; CI: confidence interval. Note: interpretation of intraclass correlation coefficient (ICC): 0.00–0.19 = very weak; 0.20–0.39 = weak; 0.40–0.59 = moderate; 0.60–0.79 = high; 0.80–1.00 = very high.

Table 2: Descriptive data for the test sample ($n = 12{,}190$).

Item	Median (IQR)
Pain	50 (30–70)
Work limitation	50 (20–70)
Leisure limitation	70 (50–90)
Sleep limitation	30 (10–60)
Independence limitation	20 (0–50)
Subjective body part value	50 (30–70)
Sick leave	10 (10–10)

IQR: interquartile range.

In the subgroup of the test study population ($n = 302$), the criterion validity for pain was borderline acceptable (correlation of the item of pain between the SOMD and WOMAC of the hip: rho = 0.62) (Figure 1).

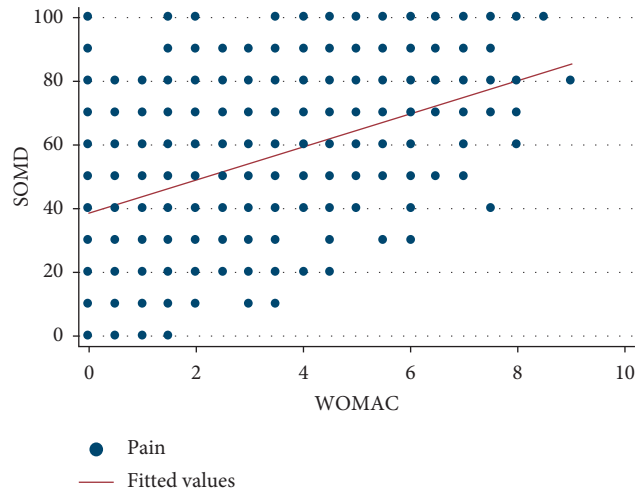

Figure 1: Item-rest correlations of all ten items of the test study population ($n = 302$).

4. Discussion

This is the first study reporting on the validation of the SOMD. Our findings show that the SOMD is highly reliable (i.e., very high test-retest reliability and internal consistency) but has borderline acceptable validity (i.e., correlation of the item of pain in the SOMD and WOMAC of the hip) and improvable clinical implementation (return rates).

Overall, the SOMD appears as a valid measurement tool for the quantification of a joint-specific HRQoL. However, several modifications are necessary for revised future versions. In general, improvement of reliability can be achieved with item homogeneity (i.e., identical minimum and maximum of different items) and item selectivity (i.e., substantial differences in minimum and maximum of a single item). For improved internal and external quality control of surgical indications and outcomes, economic analyses, and research, a highly validated questionnaire, such as the EuroQol 5-Items (EQ-5D) [10], should be integrated into the SOMD to allow calculation of a single index value and quality-adjusted life years (QALYs). The EQ-5D is a standardized, generic, self-reported, fast, generalized, and validated questionnaire. It is available in 120 languages, highly reliable and valid, and commonly used to measure HRQoL. The first version with three levels was designed in 1990 and its current form with five levels (no, slight, moderate, severe, and extreme) has been used since 2009. It allows the calculation of 3,125 (5^5) different HRQoL states. Using crosswalk links, a single index value and QALYs (0 = dead to 1 = perfect health) can be calculated [11]. Using population value sets, this permits utility analyses. Furthermore, questions from the validated Core Outcome Measures Index (COMI) questionnaire [12] could be added to quantify the joint-specific HRQoL and level of pain. Additionally, survey reports can vary according to the educational level [13], which should be quantified by an additional question. We have proposed a revised version of the SOMD, which was drafted by a subcommittee of the Swiss Orthopaedics Panel of Quality and Methodology together with the general manager of the Swiss Implant Register (SIRIS) [14] and has been approved by Swiss Orthopaedics board of directors [15].

5. Conclusion

In conclusion, this is the first report about the validation of the SOMD. A relatively high reliability (test-retest and internal consistency), borderline acceptable (criterion) validity for the item of pain and improvable clinical implementation was observed. The analysis serves as the basis for a structured modification of an improved version of the SOMD.

Data Availability

All data analyzed during this study are included in this article.

Disclosure

The subcommittee of the Swiss Orthopaedics Panel of Quality and Methodology of Swiss Orthopaedics consisted of the authors MF and CD, as well as Anne Lübbeke, MD, M.Sc., DSc, and Anne Mannion, MD, PhD. The general manager of the Swiss Implant Register (SIRIS) is Andreas Mischler. The abstract was presented as an oral presentation at the Annual Conference of the Swiss Orthopaedics, Montreux, Switzerland, June 06–08, 2018.

Conflicts of Interest

The authors declare that they have no competing interests.

Authors' Contributions

TJ was responsible for conception and design, acquisition of data, analysis and interpretation of data, and drafting the manuscript. CD interpretated the data. UM was responsible for acquisition of data and interpretation of data. MF was responsible for the idea, conception and design, ethics, interpretation of data, and supervision of the study. All authors revised and approved the final version of the manuscript.

References

[1] I. Daig and A. Lehmann, "Procedures to measure quality of life," *Zeitschrift für Medizinische Psychologie*, vol. 16, pp. 5–23, 2007.

[2] N. Bellamy, *Osteoarthritis—an Evaluative Index for Clinical Trials*, McMaster University, Hamilton, Canada, 1982.

[3] N. Bellamy, W. W. Buchanan, C. H. Goldsmith, J. Campbell, and L. W. Stitt, "Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee," *The Journal of Rheumatology*, vol. 15, no. 12, pp. 1833–1840, 1988.

[4] F. Wolfe and S. X. Kong, "Rasch analysis of the Western Ontario MacMaster questionnaire (WOMAC) in 2205 patients with osteoarthritis, rheumatoid arthritis, and fibromyalgia," *Annals of the Rheumatic Diseases*, vol. 58, no. 9, pp. 563–568, 1999.

[5] F. Angst, T. Ewert, S. Lehmann, A. Aeschlimann, and G. Stucki, "The factor subdimensions of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) help to specify hip and knee osteoarthritis. a prospective evaluation and validation study," *The Journal of Rheumatology*, vol. 32, no. 7, pp. 1324–1330, 2005.

[6] F. Angst, A. Aeschlimann, and G. Stucki, "Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities," *Arthritis & Rheumatism*, vol. 45, no. 4, pp. 384–391, 2001 Aug.

[7] N. J. Collins, D. Misra, D. T. Felson, K. M. Crossley, and E. M. Roos, "Measures of knee function: international knee documentation committee (IKDC) subjective knee evaluation form, knee injury and osteoarthritis outcome score (KOOS), knee injury and osteoarthritis outcome score physical function short form (KOOS-PS), knee outcome survey activities of daily living scale (KOS-ADL), lysholm knee scoring scale, Oxford knee score (OKS), western Ontario and McMaster Universities osteoarthritis index (WOMAC), activity rating scale (ARS), and tegner activity score (TAS)," *Arthritis Care & Research*, vol. 63, no. S11, pp. S208–S228, 2011.

[8] T. Ngo, M. Stupar, P. Côté, E. Boyle, and H. Shearer, "A study of the test-retest reliability of the self-perceived general recovery and self-perceived change in neck pain questions in patients with recent whiplash-associated disorders," *European Spine Journal*, vol. 19, no. 6, pp. 957–962, 2010.

[9] J. C. Nunnally and I. H. Bernstein, *Psychometric Theory*, McGraw–Hill, New York, NY, USA, 3rd edition, 1994.

[10] The EuroQol Group, "EuroQol-a new facility for the measurement of health-related quality of life," *Health Policy*, vol. 16, no. 3, pp. 199–208, 1990.

[11] R. Rabin and F. d. Charro, "EQ-SD: a measure of health status from the EuroQol Group," *Annals of Medicine*, vol. 33, no. 5, pp. 337–343, 2001.

[12] A. F. Mannion, F. Porchet, F. S. Kleinstück et al., "The quality of spine surgery from the patient's perspective. Part 1: the Core Outcome Measures Index in clinical practice," *European Spine Journal*, vol. 18, no. S3, pp. 367–373, 2009.

[13] M. E. Greene, O. Rolfson, S. Nemes, M. Gordon, H. Malchau, and G. Garellick, "Education attainment is associated with patient-reported outcomes: findings from the Swedish Hip Arthroplasty Register," *Clinical Orthopaedics and Related Research*, vol. 472, no. 6, pp. 1868–1876, 2014.

[14] Swiss National Joint Registry, "SIRIS report 2012-2016: annual report of the swiss national joint registry, hip and knee," Technical Report, pp. 23–38, University of Bern, Bern, Switzerland, 2016.

[15] Foundation for quality assurance in implant surgery—SIRIS, 2020, https://www.siris-implant.ch/en/Downloads?category=32.